Revisiting Resource Pooling: The Case for In-network Resource Sharing

Lorenzo Saino

Department of Electronic and Electrical Engineering
University College London

Joint work with Ioannis Psaras and George Pavlou
Outline

• Background
  – Resource pooling
  – Information Centric Networking

• In-network resource pooling
  – Main concepts
  – High level operation

• Early results

• Summary
Resource Pooling
Resource Pooling

• Well known principle in systems design supporting efficient utilization of resources under variable user demands.
Resource Pooling

• Well known principle in systems design supporting efficient utilization of resources under variable user demands.

• Resource pooling deeply routed in Internet architecture:
  – Packet switching enables pooling of link capacities and routers processing power
  – Buffers enable pooling of link capacity at adjacent time periods
  – MPLS traffic engineering and ECMP enable pooling of multiple paths
Resource Pooling

• Well known principle in systems design supporting efficient utilization of resources under variable user demands.

• Resource pooling deeply routed in Internet architecture:
  – Packet switching enables pooling of link capacities and routers processing power
  – Buffers enable pooling of link capacity at adjacent time periods
  – MPLS traffic engineering and ECMP enable pooling of multiple paths

• Pooled resources:
  – Router processing power
  – Links
  – Buffers
  – Paths
Efficiently Pooling End-to-end Paths
Efficiently Pooling End-to-end Paths

- Multipath TCP has been recently proposed to efficiently pool end-to-end paths
Efficiently Pooling End-to-end Paths

• Multipath TCP has been recently proposed to efficiently pool end-to-end paths
• Multiple simultaneous connections are opened between two communicating hosts over different paths
Efficiently Pooling End-to-end Paths

- Multipath TCP has been recently proposed to efficiently pool end-to-end paths
- Multiple simultaneous connections are opened between two communicating hosts over different paths
- Load is dynamically shifted among each path based on available bandwidth
Efficiently Pooling End-to-end Paths

- Multipath TCP has been recently proposed to efficiently pool end-to-end paths
- Multiple simultaneous connections are opened between two communicating hosts over different paths
- Load is dynamically shifted among each path based on available bandwidth
- Assumes that at least one host is multihomed
Efficiently Pooling End-to-end Paths

- Multipath TCP has been recently proposed to efficiently pool end-to-end paths
- Multiple simultaneous connections are opened between two communicating hosts over different paths
- Load is dynamically shifted among each path based on available bandwidth
- Assumes that at least one host is multihomed
- More reactive and fine-grained control than MPLS traffic engineering and ECMP
The long discussion on TCP
The long discussion on TCP

- TCP addresses uncertainty using the packet conservation principle and by (proactively) suppressing demand
The long discussion on TCP

- TCP addresses uncertainty using the packet conservation principle and by (proactively) suppressing demand
- TCP is moving traffic as fast as the path’s slowest link
The long discussion on TCP

- TCP addresses uncertainty using the packet conservation principle and by (proactively) suppressing demand.
- TCP is moving traffic as fast as the path’s slowest link.
- End-points have to speculate on the resources available along the end-to-end path.

![Diagram showing resource management](i) e2e Resource Management
Information Centric Networking (ICN)
Information Centric Networking (ICN)

- ICN paradigm aims at rethinking Internet architecture having in mind content distribution as the main use case
Information Centric Networking (ICN)

• ICN paradigm aims at rethinking Internet architecture having in mind content distribution as the main use case
• CCN/NDN [CoNEXT’09] is the most prominent architecture
Information Centric Networking (ICN)

- ICN paradigm aims at rethinking Internet architecture having in mind content distribution as the main use case
- CCN/NDN [CoNEXT’09] is the most prominent architecture
- Main principles:
  - Naming contents instead of hosts
  - Receiver-driven request-response mode of operation
  - Securing content, not channel
  - Ubiquitous packet caches on routers
ICN operation

www.example.com/video1.jpg/1
ICN operation

www.example.com/video1.jpg/1
ICN operation

www.example.com/video1.jpg/1
ICN operation

www.example.com/video1.jpg/1
ICN operation

www.example.com/video1.jpg/1
ICN operation

www.example.com/video1.jpg/1
ICN operation

www.example.com

R1

C

S

www.example.com/video1.jpg/1
ICN operation

www.example.com/video1.jpg/1
ICN operation

www.example.com

www.example.com/video1.jpg/1
Transport layer implications of ICN
Transport layer implications of ICN

- Request and Data paths are symmetric
- Instead of the “data-ACK” model of TCP, in ICN we have a “request-data” model
Transport layer implications of ICN

- Request and Data paths are symmetric
- Instead of the “data-ACK” model of TCP, in ICN we have a “request-data” model

- Receivers (instead of senders) regulate the traffic that is pushed in the network
- Based on requests forwarded, each forwarding entity knows how much traffic to expect within one RTT.
In-network caches as resources

• Network caches have been used for *resource optimization* by storing popular contents, possibly for long time
  – Reduce latency, load on origin servers and bandwidth utilization
• Overlay caching:
  – Put caches in “strategic” places and redirect (HTTP) requests to those caches
In-network caches as resources

• Network caches have been used for resource optimization by storing popular contents, possibly for long time
  – Reduce latency, load on origin servers and bandwidth utilization

• Overlay caching:
  – Put caches in “strategic” places and redirect (HTTP) requests to those caches

• Packet caching:
  – Individually named and self-identifiable packets/chunks allow for in-network storage!
  – Put caches in every router and serve network-layer requests for named chunks from caches on the path
In-network caches as resources

• Network caches have been used for *resource optimization* by storing popular contents, possibly for long time
  – Reduce latency, load on origin servers and bandwidth utilization
• Overlay caching:
  – Put caches in “strategic” places and redirect (HTTP) requests to those caches
• Packet caching:
  – Individually named and self-identifiable packets/chunks allow for in-network storage!
  – Put caches in every router and serve network-layer requests for named chunks from caches on the path
• We use in-network caching for *temporary storage*
Caches and resource pooling

• The presence of ubiquitous packet caches enables more efficient usage of resources by enabling pooling of sub-paths.
• More effective than buffers
Pooled resources
Pooled resources

- Links
- Switching devices
- Buffers

Packet switching
<table>
<thead>
<tr>
<th>Pooled resources</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Links</strong></td>
</tr>
<tr>
<td><strong>Switching devices</strong></td>
</tr>
<tr>
<td><strong>Buffers</strong></td>
</tr>
<tr>
<td><strong>Paths</strong></td>
</tr>
</tbody>
</table>

Packet switching

ECMP, MPLS TE, MPTCP
Pooled resources

- Links
- Switching devices
- Buffers
- Paths
- Sub-paths
- Packet caches

Packet switching
ECMP, MPLS TE, MPTCP
Our proposal
Proposed solution

1. Push traffic *as far in the path and as fast* as possible.
2. Once in front of the bottleneck, *store traffic temporarily* in custodian nodes/routers and deal with congestion locally.
3. Exploit all available (sub-)paths making decisions on a *hop-by-hop manner.*
Three-phase operation

• **Push-data phase** – Open-loop system
  – Receivers request for as much data as supported by their access link
  – Senders push data as far and as quickly as possible

• **Cache & Detour phase**
  – Every router monitors rate of incoming *Requests*
  – When demand is expected to exceed supply, the local router tries to find alternative paths to detour
  – In the meantime traffic in excess (if any) is cached locally

• **Backpressure phase** – Closed-loop system
  – If alternative paths do not exist or are equally congested:
    • Pace requests
    • Send notification upstream to slow down and enter closed-loop transmission
Three-phase operation

**Push-data phase** – open-loop system
- Receivers request for as much data as supported by their access link
- Senders push data as far and as quickly as possible
Three-phase operation

**Push-data phase** – open-loop system

- Receivers request for as much data as supported by their access link
- Senders push data as far and as quickly as possible
Three-phase operation

**Push-data phase** – open-loop system

- Receivers request for as much data as supported by their access link
- Senders push data as far and as quickly as possible
Three-phase operation

Cache & Detour phase

- Every router monitors rate of incoming \textit{Requests}
- When demand is expected to exceed supply, the local router tries to find alternative paths to detour
- In the meantime traffic in excess (if any) is cached locally
Three-phase operation

Cache & Detour phase

- Every router monitors rate of incoming Requests
- When demand is expected to exceed supply, the local router tries to find alternative paths to detour
- In the meantime traffic in excess (if any) is cached locally
Three-phase operation

Cache & Detour phase

- Every router monitors rate of incoming Requests
- When demand is expected to exceed supply, the local router tries to find alternative paths to detour
- In the meantime traffic in excess (if any) is cached locally
Three-phase operation

Cache & Detour phase

- Every router monitors rate of incoming *Requests*
- When demand is expected to exceed supply, the local router tries to find alternative paths to detour
- In the meantime traffic in excess (if any) is cached locally
Three-phase operation

**Backpressure phase** – Closed-loop system

If alternative paths do not exist or are equally congested:

- Pace requests
- Send notification upstream to slow down and enter closed-loop transmission
Three-phase operation

**Backpressure phase** – Closed-loop system

If alternative paths do not exist or are equally congested:

- Pace requests
- Send notification upstream to slow down and enter closed-loop transmission
Feasibility
Feasibility

- Recent work [CuckooSwitch, CoNEXT’13], [MICA, NSDI’14], [Caesar, ANCS’14] shows the feasibility of per-packet read/write operations in DRAM at several 10s Gbps on off-the-shelf hardware.
Feasibility

- Recent work [CuckooSwitch, CoNEXT’13], [MICA, NSDI’14], [Caesar, ANCS’14] shows the feasibility of per-packet read/write operations in DRAM at several 10s Gbps on off-the-shelf hardware.
- DRAM cost is steadily decreasing. We’ll soon have TBs of DRAM available on commodity servers.
Feasibility

• Recent work [CuckooSwitch, CoNEXT’13], [MICA, NSDI’14], [Caesar, ANCS’14] shows the feasibility of per-packet read/write operations in DRAM at several 10s Gbps on off-the-shelf hardware.

• DRAM cost is steadily decreasing. We’ll soon have TBs of DRAM available on commodity servers.

• Ongoing work suggests also Flash-based packet caches could be a viable solution.
### Availability of detour paths

<table>
<thead>
<tr>
<th>ISP</th>
<th>1 hop</th>
<th>2 hops</th>
<th>3+ hops</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exodus (US)</td>
<td>49.77%</td>
<td>35.48%</td>
<td>6.68%</td>
<td>8.06%</td>
</tr>
<tr>
<td>VSNL (IN)</td>
<td>25.00%</td>
<td>33.33%</td>
<td>0.00%</td>
<td>41.67%</td>
</tr>
<tr>
<td>Level 3</td>
<td>92.22%</td>
<td>6.55%</td>
<td>0.68%</td>
<td>0.55%</td>
</tr>
<tr>
<td>Sprint (US)</td>
<td>56.66%</td>
<td>37.08%</td>
<td>1.81%</td>
<td>4.45%</td>
</tr>
<tr>
<td>AT&amp;T (US)</td>
<td>34.84%</td>
<td>61.69%</td>
<td>0.72%</td>
<td>2.74%</td>
</tr>
<tr>
<td>EBONE (EU)</td>
<td>50.66%</td>
<td>36.22%</td>
<td>6.30%</td>
<td>6.82%</td>
</tr>
<tr>
<td>Telstra (AUS)</td>
<td>70.05%</td>
<td>10.42%</td>
<td>1.06%</td>
<td>18.47%</td>
</tr>
<tr>
<td>Tiscali (EU)</td>
<td>24.50%</td>
<td>39.85%</td>
<td>10.15%</td>
<td>25.50%</td>
</tr>
<tr>
<td>Verio (US)</td>
<td>71.50%</td>
<td>17.09%</td>
<td>1.74%</td>
<td>9.68%</td>
</tr>
<tr>
<td><strong>Average</strong></td>
<td>52.80%</td>
<td>30.86%</td>
<td>3.24%</td>
<td>13.10%</td>
</tr>
</tbody>
</table>
Some (very initial) results

![Bar chart showing network throughput for Telstra, Exodus, and Tiscali. The chart compares SP, ECMP, and INRP.](image-url)
Summary and open issues

• Information-Centric Networking:
  – Lots of attention lately
  – Requires investment and effort
  – Worth doing, but need to get the full set of advantages

• There is an opportunity to deal with congestion control at the network layer
Summary and open issues

- Information-Centric Networking:
  - Lots of attention lately
  - Requires investment and effort
  - Worth doing, but need to get the full set of advantages

- There is an opportunity to deal with congestion control at the network layer

- Open Issues:
  - How do you know detour paths are not congested
  - How will this co-exist with traditional TCP flows?
  - Out of order delivery
  - Flows swapping between original and detour paths